ELSEVIER

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Discrete Applied Mathematics 128 (2003) 275–292

DISCRETE APPLIED MATHEMATICS

www.elsevier.com/locate/dam

The average dimension of the hull of cyclic codes

Gintaras Skersys

Matematikos ir Informatikos Fakultetas, Vilniaus Universitetas, Naugarduko 24, 2006 Vilnius, Lithuania

Received 1 March 2001; received in revised form 26 October 2001; accepted 8 April 2002

Abstract

We study $E_q(n)$, the average dimension of the hull of error-correcting block cyclic codes of a given length n over a given finite field $\mathbf{F}_q$, where the hull of a code is its intersection with its dual code. We derive an expression of $E_q(n)$ which handles well. Using this expression, we prove that either $E_q(n)$ is zero (if, and only if, $n \in \mathcal{N}_q$), or it grows at the same rate as n, when $n \notin \mathcal{N}_q$, where $\mathcal{N}_q$ is the set of positive divisors of the integers of the form $q^i + 1$, $i > 0$. This permits us to show that, for almost all n, the hull of most cyclic codes of length n is "large". Moreover, we study the asymptotic behaviour of $E_q(n)/n$ as n tends to infinity.

Keywords: Error-correcting codes; Cyclic codes; Hull

1. Introduction

Recently, an algorithm for computing the permutation and automorphism groups of an error-correcting block linear code, and for determining the equivalence and the permutation equivalence of two such codes, based on the algorithms of Leon [4–6] and Sendrier [10], was presented [11–13]. This algorithm is limited by the size of the hull (the intersection of a code with its dual). That shows the necessity to study the size of the hull of linear codes. In [9], Sendrier studies the expected dimension of the hull of a random linear code. He shows that asymptotically it is a small positive constant. The present paper concerns the expected dimension of the hull of a random

E-mail address: gintaras.skersys@maf.vu.lt (G. Skersys).

0166-218X/03/$ - see front matter
PII: S0166-218X(02)00451-1

cyclic code. We show that either it is zero, or it grows at the same rate as the length of codes.

In this paper q is a power of a prime p (except in Section 2), $\mathbf{F}_q$ is the finite field of size q, n is a positive integer, r and s are determined uniquely by n as follows: $n = rp^s$, r and p are relatively prime, $s \geqslant 0$.

See [8] for basic definitions on error-correcting codes. Using standard terminology of coding theory, a *cyclic code* of length n is an ideal C in the ring $\mathbf{F}_q[X]/(X^n - 1)$. The *dimension* of C, denoted by $\dim C$, is the dimension of C considered as linear space over $\mathbf{F}_q$. The *dual code* $C^\perp$ of C is defined to be $C^\perp = \{\mathbf{u} \in \mathbf{F}_q^n \mid \mathbf{u} \cdot \mathbf{v} = 0 \text{ for all } \mathbf{v} \in C\}$, where $\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + \cdots + u_n v_n$ is the *scalar product* of vectors $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$. The hull was introduced by Assmus and Key [1]. The *hull* of a linear code C, denoted by $\mathscr{H}(C)$, is its intersection with its dual code: $\mathscr{H}(C) = C \cap C^\perp$. Let $C(n,q)$ be the set of cyclic codes of length n over $\mathbf{F}_q$. We study $E_q(n)$, the average dimension of the hull of cyclic codes of a given length n over a given finite field $\mathbf{F}_q$, i.e.

$$E_q(n) = \frac{\sum_{C \in C(n,q)} \dim \mathscr{H}(C)}{|C(n,q)|},$$

where by $|A|$ we denote the size of the set A.

Note that we can define $E_q(n)$ in terms of the probability theory. Let Y be the random variable that takes as value $\dim \mathscr{H}(C)$ when we choose at random a cyclic code C from $C(n,q)$ with uniform probability. One can easily show that the expectation $E(Y) = E_q(n)$.

Section 2 defines the set $\mathscr{N}_q$, gives its properties and introduces the $\mathscr{N}_q$-factorization of an integer. Section 3 derives an expression of $E_q(n)$ (Theorem 10). Section 4 determines when $E_q(n)$ is zero and proves that if $E_q(n)$ is not zero, it grows at the same rate as n. And finally Section 5 studies the asymptotic behavior of $E_q(n)/n$ as n tends to infinity.

2. $\mathscr{N}_q$-factorization

The set $\mathscr{N}_q$ is of great importance in the study of $E_q(n)$. In this section we show the structure of $\mathscr{N}_q$. These results allow us to define the $\mathscr{N}_q$-factorization of an integer. The $\mathscr{N}_q$-factorization allows us to find an expression of $E_q(n)$ easier to handle (Theorem 10).

Note that the results of this section are valid for any integer $q \geqslant 2$.

$\mathscr{N}_q$ is the set of positive divisors of the integers of the form $q^i + 1$, $i > 0$, i.e.

$$\mathscr{N}_q = \{l \geqslant 1 \colon \exists i = i(q,l) \geqslant 1 \text{ such that } l \mid q^i + 1\}.$$

(Recall that the vertical slash means "divides".) Note that $l \mid q^i + 1$ is equivalent to $q^i \equiv -1 \pmod{l}$. We use both notations.

The set $\mathscr{N}_q$ was studied independently by Knee and Goldman [3] and the author [13]. The approaches are different, but for our purposes the results published in [3] are almost sufficient, so we shall give them without proof and use them. Note that the

results in [3] are given for q a prime power (some of them for q a prime), but the generalization for any $q \geqslant 2$ is straightforward [13]. We begin by a simple lemma:

Lemma 1 (Knee and Goldman [3, Theorem 3]). (1) *If $l \in \mathcal{N}_q$, then l and q are relatively prime.*
(2) *If $l \in \mathcal{N}_q$ and $d > 0$ divides l, then $d \in \mathcal{N}_q$.*

Let $(l,q) = 1$, where (l,q) denotes the greatest common divisor of l and q. Denote $\mathrm{ord}_l(q)$ the multiplicative order of q modulo l, i.e. the smallest positive integer t such that $q^t \equiv 1 \pmod{l}$. Remind that if $s > 0$, then $q^s \equiv 1 \pmod{l}$ if, and only if, $\mathrm{ord}_l(q) \mid s$. In particular, $\mathrm{ord}_l(q) \mid \varphi(l)$, where $\varphi(l)$ is Euler's function, it gives the number of integers i with $1 \leqslant i \leqslant l$ that are relatively prime to l. By part 1 of Lemma 1, if $l \in \mathcal{N}_q$, $\mathrm{ord}_l(q)$ is well defined.

Lemma 2. *If $l \in \mathcal{N}_q$, $l > 2$, then $\mathrm{ord}_l(q)$ is even.*

Proof. If $l \in \mathcal{N}_q$, then there exists $i > 0$ such that $q^i \equiv -1 \pmod{l}$, and, since $l > 2$, $q^i \not\equiv 1 \pmod{l}$, so that $\mathrm{ord}_l(q) \nmid i$. But $q^{2i} \equiv (-1)^2 \equiv 1 \pmod{l}$, so that $\mathrm{ord}_l(q) \mid 2i$. Therefore, $\mathrm{ord}_l(q)$ is even. □

Now we study the structure of $\mathcal{N}_q$. It turns out that we may partition the elements of $\mathcal{N}_q$ in some natural way. For $\alpha \geqslant 0$ we denote

$$\mathcal{P}_{q,\alpha} = \{l \in \mathcal{N}_q \colon 2^\alpha \parallel \mathrm{ord}_l(q)\},$$

where $2^\alpha \parallel v$ signifies that 2^α is the greatest power of 2 dividing v.

Remark 3. By definition, $1 \in \mathcal{N}_q$. We may define $\mathrm{ord}_1(q)$ to be 1, so we consider that $1 \in \mathcal{P}_{q,0}$. If q is odd, then $2 \mid q+1$, so $2 \in \mathcal{N}_q$. Moreover, $\mathrm{ord}_2(q) = 1$, so $2 \in \mathcal{P}_{q,0}$. If q is even, $2 \notin \mathcal{N}_q$ by part 1 of Lemma 1. The set $\mathcal{P}_{q,0}$ has no more elements, because by Lemma 2, if an integer $l > 2$ belongs to $\mathcal{N}_q$, then $\mathrm{ord}_l(q)$ is even, so $l \in \mathcal{P}_{q,\alpha}$ with $\alpha \geqslant 1$. □

Clearly, for all $l \in \mathcal{N}_q$, $\mathrm{ord}_l(q)$ is finite, so that the corresponding α is determined uniquely. Since for all α, $\mathcal{P}_{q,\alpha} \neq \emptyset$ (Remark 3 and Proposition 7), it follows that $\{\mathcal{P}_{q,\alpha}\}_{\alpha \geqslant 0}$ is an (infinite) partition of $\mathcal{N}_q$. Now we give the result that expresses the membership of l in $\mathcal{N}_q$ by means of the prime factors of l.

Theorem 4. *Let $q \geqslant 2$.*

(a) *An odd integer $l = \prod_i u_i^{e_i}$ (prime factorization) belongs to $\mathcal{N}_q$ if, and only if, there exists $\alpha \geqslant 0$ such that $u_i \in \mathcal{P}_{q,\alpha}$ for all i. Also in this case $l \in \mathcal{P}_{q,\alpha}$.*
(b) *Let $\beta \geqslant 1$. Then $2^\beta \in \mathcal{N}_q$ if, and only if, 2^β divides $q+1$. Moreover, if $2^\beta \in \mathcal{N}_q$, $\beta \geqslant 2$, then $2^\beta \in \mathcal{P}_{q,1}$.*
(c) *Let q and $l > 0$ be odd. Then $2l \in \mathcal{N}_q$ if, and only if, $l \in \mathcal{N}_q$. Also in this case l and $2l$ belong to the same set $\mathcal{P}_{q,\alpha}$.*

(d) *Let* $l = 2^\beta t$, *where* t *is odd*, $\beta \geqslant 2$. *Then* $l \in \mathscr{N}_q$ *if, and only if,* $2^\beta \in \mathscr{N}_q$ *and* $t \in \mathscr{P}_{q,1}$. *Also in this case* $l \in \mathscr{P}_{q,1}$.

Proof. This theorem (except the second parts of parts (b)–(d)) is a generalized for any $q \geqslant 2$ and reformulated version of Theorems 5, 6a and 7a in [3]. This generalization is straightforward, so that we refer the reader to [3] for the proof.

It remains to prove the second parts of parts (b)–(d). We begin by part (b). Let $2^\beta \in \mathscr{N}_q$, $\beta \geqslant 2$. We shall show that $2^\beta \in \mathscr{P}_{q,1}$. By the first part of part (b), 2^β divides $q+1$, that is $q \equiv -1 \pmod{2^\beta}$. Therefore $\mathrm{ord}_{2^\beta}(q) \neq 1$ and, since $q^2 \equiv 1 \pmod{2^\beta}$, $\mathrm{ord}_{2^\beta}(q) \mid 2$, so that $\mathrm{ord}_{2^\beta}(q) = 2$ and $2^\beta \in \mathscr{P}_{q,1}$.

Now part (c). Let $l \in \mathscr{P}_{q,\alpha}$ for some $\alpha \geqslant 0$. If $\alpha = 0$, then by Remark 3 $l = 1$ and $2 \in \mathscr{P}_{q,0}$, since q is odd. Suppose now $\alpha \geqslant 1$, so that $l > 2$. Since $l \in \mathscr{N}_q$, there exists $i > 0$ such that $q^i \equiv -1 \pmod{l}$, so that $\mathrm{ord}_l(q) \nmid i$, but as above $\mathrm{ord}_l(q) \mid 2i$. But $2^\alpha \parallel \mathrm{ord}_l(q)$, therefore $2^\alpha \parallel 2i$. Moreover, since $l \mid q^i + 1$ and l and q are odd, we have $2l \mid q^i + 1$ and $q^i \equiv -1 \pmod{2l}$. Therefore, if ξ is such that $2^\xi \parallel \mathrm{ord}_{2l}(q)$, in the same manner we show that $2^\xi \parallel 2i$, so that $\xi = \alpha$ and $2l \in \mathscr{P}_{q,\alpha}$.

Finally part (d). Let $l = 2^\beta t$, t odd, $\beta \geqslant 2$. Let $l \in \mathscr{N}_q$, i.e. $t \in \mathscr{P}_{q,1}$ and $2^\beta \in \mathscr{P}_{q,1}$ (part (b)). Let i be such that $q^i \equiv -1 \pmod{t}$. As above we show that, since $2 \parallel \mathrm{ord}_t(q)$, we have $2 \parallel 2i$, so that i is odd. Moreover, by part (b), $q \equiv -1 \pmod{2^\beta}$, so that $q^i \equiv (-1)^i \equiv -1 \pmod{2^\beta}$ and therefore $q^i \equiv -1 \pmod{l}$. In the same manner we show that if ξ is such that $2^\xi \parallel \mathrm{ord}_l(q)$, we have $2^\xi \parallel 2i$, so that $\xi = 1$ and $l \in \mathscr{P}_{q,1}$. □

Corollary 5. *Let* $\gamma \geqslant 0$ *be such that* $2^\gamma \parallel q+1$. *Let* l *be a positive integer relatively prime to* q, *let* $2^\beta \parallel l$.

(1) $\mathscr{P}_{q,0} = \{1\}$, *if* q *is even*, $\mathscr{P}_{q,0} = \{1,2\}$, *otherwise*.
(2) $l \in \mathscr{P}_{q,1}$ *if, and only if, either* l *has an odd prime divisor, each odd prime divisor of* l *belongs to* $\mathscr{P}_{q,1}$ *and* $0 \leqslant \beta \leqslant \gamma$, *or* $l = 2^\beta$ *and* $2 \leqslant \beta \leqslant \gamma$.
(3) *Let* $\alpha \geqslant 2$. *Then* $l \in \mathscr{P}_{q,\alpha}$ *if, and only if,* l *has an odd prime divisor, each odd prime divisor of* l *belongs to* $\mathscr{P}_{q,\alpha}$ *and* $0 \leqslant \beta \leqslant 1$.

Proof. Part 1 follows from Remark 3.

Let $l \geqslant 1$, $l = 2^\beta t$, t odd. In order to prove parts 2 and 3, we consider five cases: (1) $t > 1$ and $\beta \geqslant 2$, (2) $t > 1$ and $\beta = 1$, (3) $t > 1$ and $\beta = 0$, (4) $t = 1$ and $\beta \geqslant 2$, (5) $t = 1$ and $0 \leqslant \beta \leqslant 1$.

(1) For part 2, the result follows from parts (d), (a) and (b) of Theorem 4. For part 3, Theorem 4(d) shows that $l \notin \mathscr{P}_{q,\alpha}$, $\alpha \geqslant 2$.
(2) For both parts, the result follows from parts (c) and (a) of Theorem 4.
(3) For both parts, the result follows from Theorem 4(a).
(4) For part 2, the result follows from Theorem 4(b). For part 3, Theorem 4(b) shows that $l \notin \mathscr{P}_{q,\alpha}$, $\alpha \geqslant 2$.
(5) For both parts, Remark 3 asserts that $l \notin \mathscr{P}_{q,\alpha}$, $\alpha \geqslant 1$. □

Corollary 6. *If an integer $l > 2$ belongs to $\mathscr{P}_{q,\alpha}$, $\alpha \geqslant 1$, then every divisor $d > 2$ of l belongs to $\mathscr{P}_{q,\alpha}$ too.*

Proof. We shall prove the corollary in the case when $\alpha \geqslant 2$. If $\alpha = 1$, the proof, based on part (2) of Corollary 5, is analogous. Since $l \in \mathscr{P}_{q,\alpha}$, l satisfies the conditions of part (3) of Corollary 5. Every divisor $d > 2$ of l satisfies these conditions too, so $d \in \mathscr{P}_{q,\alpha}$. □

Proposition 7. *Let $\alpha \geqslant 1$. Then $\mathscr{P}_{q,\alpha}$ is the set of divisors > 2 of the integers of the form $q^i + 1$, $i = 2^{\alpha-1} + 2^\alpha k$, $k \geqslant 0$.*

Proof. We begin by showing that $\operatorname{ord}_{q^i+1}(q) = 2i$ for all $i \geqslant 1$. Indeed, $q^i \equiv -1 \pmod{q^i + 1}$, so $q^{2i} \equiv (-1)^2 \equiv 1 \pmod{q^i + 1}$ and $\operatorname{ord}_{q^i+1}(q) \mid 2i$. Let d be a divisor of $2i$, $d < 2i$. Then $d \leqslant i$, so $1 < q^d \leqslant q^i < q^i + 1$, whence $q^d \not\equiv 1 \pmod{q^i + 1}$ and $d \neq \operatorname{ord}_{q^i+1}(q)$. Thus, $\operatorname{ord}_{q^i+1}(q) = 2i$.

Let l be a divisor > 2 of an integer of the form $q^i + 1$, $i = 2^{\alpha-1} + 2^\alpha k$, $k \geqslant 0$. Then $\operatorname{ord}_{q^i+1}(q) = 2i = 2^\alpha + 2^{\alpha+1}k$, so that $2^\alpha \parallel \operatorname{ord}_{q^i+1}(q)$ and $q^i + 1 \in \mathscr{P}_{q,\alpha}$. By Corollary 6, l also belongs to $\mathscr{P}_{q,\alpha}$.

Conversely, let $l \in \mathscr{P}_{q,\alpha}$. By Remark 3, $l > 2$. Since $l \in \mathscr{N}_q$, there exists $i \geqslant 1$ such that l divides $q^i + 1$. By Corollary 6, $q^i + 1 \in \mathscr{P}_{q,\alpha}$ too, so that $2^\alpha \parallel \operatorname{ord}_{q^i+1}(q) = 2i$. Therefore, $2^{\alpha-1} \parallel i$ and $i = 2^{\alpha-1}(1 + 2k) = 2^{\alpha-1} + 2^\alpha k$, $k \geqslant 0$. □

See [3,13] for more properties of the set $\mathscr{N}_q$.

Now we introduce the $\mathscr{N}_q$-factorization of an integer. Let l be a positive integer relatively prime to q. Let $l = 2^\beta u_1^{e_1} \cdots u_k^{e_k}$ be the prime factorization of l, $k \geqslant 0$, $\beta \geqslant 0$, $e_i > 0$, u_i an odd prime for all i, $1 \leqslant i \leqslant k$. We may partition the set of indices $\{1, \ldots, k\}$ into (possibly empty) subsets $K', K_1, K_2, \ldots$ as follows: i, $1 \leqslant i \leqslant k$, belongs to K' if, and only if, u_i does not belong to $\mathscr{N}_q$, and belongs to K_α, $\alpha \geqslant 1$, if, and only if, u_i belongs to $\mathscr{P}_{q,\alpha}$. Let us denote $d_\alpha = \prod_{i \in K_\alpha} u_i^{e_i}$, $\alpha \geqslant 1$, and $d' = \prod_{i \in K'} u_i^{e_i}$ (with the convention $\prod_{i \in \emptyset} u_i^{e_i} = 1$). Then we get the following factorization of l: $l = 2^\beta d' d_1 d_2 d_3 \cdots$. We call it the *$\mathscr{N}_q$-factorization* of l. Note that $d_\alpha = 1$ for all $\alpha \geqslant 1$ except a finite set of them. Note also that, according to Corollary 5, for all $\alpha \geqslant 1$, if $d_\alpha > 1$, then $d_\alpha \in \mathscr{P}_{q,\alpha}$. Denote by $\mathscr{P}'_q$ the set of odd integers > 1, relatively prime to q, such that none of their prime divisors belongs to $\mathscr{N}_q$. Then if $d' > 1$, $d' \in \mathscr{P}'_q$.

As the $\mathscr{N}_q$-factorization is unique, for any positive integer l the quantities β, d' and d_α, $\alpha \geqslant 1$, are defined and unique. So we can consider them as functions of l and write $\beta(l)$, $d'(l)$, $d_\alpha(l)$, $\alpha \geqslant 1$.

We conclude this section with two lemmas we shall need in the proof of Theorem 25.

Lemma 8. *Let $\alpha \geqslant 1$. If $l \in \mathscr{P}_{q,\alpha}$, then $l \geqslant 2^\alpha + 1$.*

Proof. Let $l \in \mathscr{P}_{q,\alpha}$, $\alpha \geqslant 1$. By Remark 3, $l > 2$. Then $2^\alpha \parallel \operatorname{ord}_l(q) \mid \varphi(l) \leqslant l - 1$, so $l \geqslant 2^\alpha + 1$. □

Lemma 9. *Let $\gamma \geqslant 0$ be such that $2^\gamma \parallel q + 1$. Let $l \geqslant 2$, let l and q be relatively prime, let $l = 2^\beta d' d_1 d_2 d_3 \cdots$ be the $\mathscr{N}_q$-factorization of l. If $l \notin \mathscr{N}_q$, then at least*

one of following conditions is valid: (1) $\beta > \gamma$, (2) $d' > 1$, (3) $\beta \geqslant 2$ *and* $d_\alpha > 1$ *for an integer* $\alpha \geqslant 2$, (4) $d_{\alpha_1} > 1$ *and* $d_{\alpha_2} > 1$ *for two distinct* $\alpha_1 \geqslant 1$ *and* $\alpha_2 \geqslant 1$.

Proof. We prove by contradiction, that is, in fact we prove that if (1) $\beta \leqslant \gamma$, (2) $d' = 1$, (3) $\beta \leqslant 1$ or $d_\alpha = 1$ for all $\alpha \geqslant 2$, and (4) $d_\alpha > 1$ for at most one $\alpha \geqslant 1$, then $l \in \mathscr{N}_q$.

Assume that conditions (1)–(4) are valid. From condition (2) we get that $l = 2^\beta d_1 d_2 d_3 \cdots$. From condition (4) we get that either (a) $d_\alpha = 1$ for all $\alpha \geqslant 1$, or (b) $d_1 > 1$ and $d_\alpha = 1$ for all $\alpha \geqslant 2$, or (c) $d_1 = 1$ and $d_\alpha > 1$ for exactly one $\alpha \geqslant 2$.

In the case (a), $l = 2^\beta$. By condition (1), $\beta \leqslant \gamma$. If $0 \leqslant \beta \leqslant 1$, by part 1 of Corollary 5, $l \in \mathscr{N}_q$. If $2 \leqslant \beta \leqslant \gamma$, $l \in \mathscr{N}_q$ by part 2 of the same corollary. In the case (b), $l = 2^\beta d_1$. Then $\beta \leqslant \gamma$ by condition (1), and $l \in \mathscr{N}_q$ by part 2 of Corollary 5. In the case (c), $l = 2^\beta d_\alpha$, $\alpha \geqslant 2$. Then $\beta \leqslant 1$ by condition (3), and $l \in \mathscr{N}_q$ by part 3 of Corollary 5. So in all cases $l \in \mathscr{N}_q$. □

3. Expression of $E_q(n)$

In this section we give and prove an expression of $E_q(n)$ (Theorem 10). This expression allows us to find bounds on $E_q(n)$ (Theorem 25) and to study the asymptotic behavior of $E_q(n)$ (Theorems 26 and 27). This whole section is devoted to the proof of the following theorem:

Theorem 10. *Let* q *be a power of a prime* p, *let* $n \geqslant 1$. *Let* $n = rp^s$, $s \geqslant 0$, r *and* p *relatively prime. Then the average dimension* $E_q(n)$ *of the hull of cyclic codes of length* n *over the finite field* $\mathbf{F}_q$ *is equal to*

$$E_q(n) = n\left(\frac{1}{3} - \frac{1}{6(p^s+1)}\right) - B_{r,q}\left(\frac{p^s+1}{12} + \frac{2-3\delta_{p^s}}{12(p^s+1)}\right), \tag{1}$$

where $\delta_{p^s} = 1$, *if* p^s *is even, and* $\delta_{p^s} = 0$, *otherwise*,

$$B_{r,q} = 2^{\min(\gamma,\beta)} d_1 + 2^{\min(\beta,1)} \sum_{\alpha \geqslant 2} (d_\alpha - 1), \tag{2}$$

$r = 2^\beta d' d_1 d_2 d_3 \cdots$ *is the* $\mathscr{N}_q$*-factorization of* r, *and* γ *is such that* $2^\gamma \| q+1$.

Note that only a finite number of terms of the sum in the expression of $B_{r,q}$ are not zeros.

We give a corollary the proof of which is straightforward.

Corollary 11. (1) $E_q(n) < n/3$.
(2) $E_q(r) = (r - B_{r,q})/4$.
(3) $E_q(r) < r/4$.

In order to prove this theorem, we express the generator polynomial of the hull of a cyclic code by means of the generator polynomial of the cyclic code (Proposition 20). For this we first recall some results on the polynomials over $\mathbf{F}_q$ (Section 3.1). Then,

using Proposition 20, we prove Eq. (1) with another expression of $B_{r,q}$ (Proposition 22), and we show that these two expressions of $B_{r,q}$ are equal (Proposition 24), using Propositions 15–17 and $\mathcal{N}_q$-factorization.

3.1. Polynomials

In this section, we give some properties of reciprocal and quasi-self-reciprocal (qsr) polynomials, factorization of X^n-1 and order of polynomials, needed in the following sections.

Take the complete factorization of X^r-1 over $\mathbf{F}_q$,

$$X^r-1=f_1f_2\cdots f_t, \tag{3}$$

where $f_1,f_2,\ldots,f_t$ are monic irreducible polynomials over $\mathbf{F}_q$. For our purposes, the exact nature of these polynomials f_i is not important. It will suffice to know two facts about them: they all are distinct [8, Chapter 7, Section 5], and Proposition 13. Then

$$X^n-1=(X^r-1)^{p^s}=\prod_{i=1}^{t}f_i^{p^s} \tag{4}$$

is the complete factorization of X^n-1 over $\mathbf{F}_q$.

Let $f(X)\in\mathbf{F}_q[X]$ be a polynomial such that $f(0)\neq 0$. Let us denote

$$\bar{f}(X)=f(0)^{-1}X^{\deg f}f(1/X), \tag{5}$$

where $\deg f$ is the degree of f. The polynomial $f^*(X)=X^{\deg f}f(1/X)$ is the *reciprocal polynomial* of f, and we multiply by the constant $f(0)^{-1}$ in order to get a monic polynomial. Using Eq. (5), one can easily prove the following lemma:

Lemma 12. (1) *If $f(X)$ is a monic polynomial, then $\bar{\bar{f}}=f$.*
(2) *If $f(X)$ and $g(X)$ are two polynomials, $\overline{fg}=\bar{f}\bar{g}$.*

The second part easily generalizes to any number of polynomials. Since

$$f_1f_2\cdots f_t=X^r-1=\overline{X^r-1}=\overline{f_1}\,\overline{f_2}\cdots\overline{f_t}$$

and since the factorization is unique and $\bar{f}$ is irreducible when f is irreducible (by Lemma 12), we have the following property.

Proposition 13. *For every f_i there exists f_j such that $\overline{f_i}=f_j$.*

So that for some polynomials f_i from Eq. (3), $\overline{f_i}=f_i$, and other polynomials come in pairs f_i, f_j, where $\overline{f_i}=f_j$ and $\overline{f_j}=f_i$.

Following [3], we call a polynomial $f(X)\in\mathbf{F}_q[X]$ qsr if its reciprocal polynomial $f^*(X)=\pm f(X)$.

Proposition 14. *A monic polynomial f is qsr if, and only if, $\bar{f} = f$.*

Proof. If f is monic qsr polynomial, by comparing the constant terms we show that if $f^*(X) = f(X)$, then $f(0) = 1$, and if $f^*(X) = -f(X)$, then $f(0) = -1$, so that $f^*(X) = f(0)f(X)$ and $\bar{f}(X) = f(0)^{-1}f^*(X) = f(X)$. Conversely, if $f = \bar{f}$, f monic, in the same manner we show that $f(0) = f(0)^{-1}$, so that $f(0) = \pm 1$. Therefore, $f(X) = \bar{f}(X) = f(0)^{-1}f^*(X) = \pm f^*(X)$, so that f is qsr. □

We recall some properties of polynomials (see Lidl and Niederreiter [7]) let $f(X) \in \mathbf{F}_q[X]$ be a polynomial such that $f(0) \neq 0$. Then the least positive integer e for which $f(X)$ divides $X^e - 1$ is called the *order* of f and is denoted by $\mathrm{ord}(f)$.

Proposition 15. *Let $e > 0$. Then the polynomial $f \in \mathbf{F}_q[X]$ with $f(0) \neq 0$ divides $X^e - 1$ if, and only if,* $\mathrm{ord}(f)$ *divides e.*

Since every polynomial f_i from Eq. (3) divides $X^r - 1$, $\mathrm{ord}(f_i)$ divides r. For every divisor d of r, we collect together the polynomials f_i of the same order d. Their product $Q_d(X) = \prod_{i:\mathrm{ord}(f_i)=d} f_i$ is called the dth *cyclotomic polynomial* over $\mathbf{F}_q$. We obtain that $X^r - 1 = \prod_{d|r} Q_d(X)$, where the product is extended over all positive divisors d of r.

Proposition 16. $\deg Q_d(X) = \varphi(d)$, *where φ is Euler's function.*

Proposition 17 (Knee and Goldman [3, Theorem 1]). *If a polynomial $f(X) \neq aX$ over $\mathbf{F}_q$ is irreducible, then $f(X)$ is qsr if, and only if,* $\mathrm{ord}(f) \in \mathcal{N}_q$.

3.2. Cyclic codes, their duals and hulls

Using standard terminology of coding theory, a *cyclic code* of length n is an ideal C in the ring $\mathbf{F}_q[X]/(X^n - 1)$, generated by a monic factor $g_C(X)$ of $X^n - 1 = (X^r - 1)^{p^s}$. Also, every monic divisor of $X^n - 1$ generates a distinct ideal. The polynomial $g_C(X)$ is called the *generator polynomial* of C. By Eq. (4), every monic divisor $g_C(X)$ of $X^n - 1$ can be expressed as

$$g_C(X) = \prod_{i=1}^{t} f_i^{\varepsilon_i}, \quad \text{where} \quad 0 \leqslant \varepsilon_i \leqslant p^s \text{ for all } i. \tag{6}$$

Let $C(n,q)$ be the set of cyclic codes of length n over $\mathbf{F}_q$. We have thus showed the following.

Proposition 18. *There exists an one-to-one correspondence between $C(n,q)$ and the set $\{(\varepsilon_1,\ldots,\varepsilon_t) \mid 0 \leqslant \varepsilon_i \leqslant p^s$ for all $i\}$: a cyclic code $C \in C(n,q)$ generated by $g_C(X)$ corresponds to t-tuple $(\varepsilon_1,\ldots,\varepsilon_t)$ such that $g_C(X) = \prod_{i=1}^{t} f_i^{\varepsilon_i}$.*

The dual code $C^{\perp}$ of a cyclic code C is also cyclic, and its generator polynomial $g_{C^{\perp}}(X) = \bar{h}(X)$, where $h(X) = (X^n - 1)/g_C(X)$ [8, Theorem 4 of Chapter 7] (see Eq. (5)) for the definition of $\bar{h}(X)$). By Proposition 13, for every f_i from Eq. (3) there exists f_j such that $\overline{f_i} = f_j$. Let $\bar{\varepsilon}_i$, $1 \leqslant i \leqslant t$, be the multiplicity of $\overline{f_i}$ in $g_C(X)$, that is, $\bar{\varepsilon}_i = \varepsilon_j$. Then we have the following result.

Proposition 19.

$$g_{C^{\perp}}(X) = \prod_{i=1}^{t} f_i^{p^s} - \bar{\varepsilon}_i.$$

Proof. Since $X^n - 1 = \prod_{i=1}^{t} f_i^{p^s}$, we have $h(X) = (X^n - 1)/g_C(X) = \prod_{i=1}^{t} f_i^{p^s - \varepsilon_i}$. By Lemma 12,

$$\bar{h}(X) = \prod_{i=1}^{t} \overline{f_i^{p^s - \varepsilon_i}} = \prod_{i=1}^{t} \overline{f_i}^{p^s - \varepsilon_i} = \prod_{i=1}^{t} f_i^{p^s - \bar{\varepsilon}_i}. \qquad \square$$

One can easily show that the hull $\mathscr{H}(C) = C \cap C^{\perp}$ of a cyclic code C is also cyclic and its generator polynomial $g_{\mathscr{H}(C)}(X) = \mathrm{lcm}(g_C(X), g_{C^{\perp}}(X))$, the least common multiple of $g_C(X)$ and $g_{C^{\perp}}(X)$. The following result is then an easy consequence of Eq. (6) and Proposition 19.

Proposition 20. $g_{\mathscr{H}(C)}(X) = \prod_{i=1}^{t} f_i^{\max(\varepsilon_i,\, p^s - \bar{\varepsilon}_i)}$.

And finally we recall a property of cyclic codes that we shall need in the proof of Proposition 22.

Proposition 21. *Let C be a cyclic code of length n over $\mathbf{F}_q$ generated by $g_C(X)$. Then* $\dim C = n - \deg g_C(X)$.

3.3. *Proof of Theorem 10*

Let J be the set of such i that f_i from Eq. (3) is qsr. We recall that f_i is qsr if, and only if, $\overline{f_i} = f_i$ (Proposition 14). We denote

$$B_{r,q} = \sum_{i \in J} \deg f_i \tag{7}$$

(in Proposition 24 we show that this definition is equivalent to that of Eq. (2)). Now we express $E_q(n)$ as a function of $B_{r,q}$. Let $\delta_v = 1$, if v is even, and $\delta_v = 0$, otherwise.

Proposition 22.

$$E_q(n) = n\left(\frac{1}{3} - \frac{1}{6(p^s+1)}\right) - B_{r,q}\left(\frac{p^s+1}{12} + \frac{2 - 3\delta_{p^s}}{12(p^s+1)}\right).$$

Proof. Let Y be the random variable defined in the introduction. Then $E_q(n)=E(Y)$, the expectation of Y. By Proposition 18, there exists an one-to-one correspondence between $C(n,q)$, the set of cyclic codes of length n over $\mathbf{F}_q$, and the set $\{(\varepsilon_1,\ldots,\varepsilon_t)\,|\,0\leqslant\varepsilon_i\leqslant p^s$ for all $i\}$. It is not difficult to show that choosing a cyclic code C from $C(n,q)$ with uniform probability $1/|\,C(n,q)|$ amounts to choosing ε_i, $0\leqslant\varepsilon_i\leqslant p^s$, independently for all i, $1\leqslant i\leqslant t$, with uniform probabilities $1/(p^s+1)$. By Propositions 21 and 20, we get

$$Y=\dim\mathscr{H}(C)=n-\deg g_{\mathscr{H}(C)}(X)=n-\sum_{i=1}^{t}(\deg f_i)\max(\varepsilon_i,\,p^s-\bar{\varepsilon}_i).$$

Denote $\beta_i=\max(\varepsilon_i,\,p^s-\bar{\varepsilon}_i)$. Then

$$E_q(n)=E(Y)=n-\sum_{i=1}^{t}(\deg f_i)E(\beta_i).$$

In order to find $E(\beta_i)$, we use the known probabilities $\Pr\{\varepsilon_i=j\}=1/(p^s+1)$ for all j, $0\leqslant j\leqslant p^s$, and all i, $1\leqslant i\leqslant t$.

- Let i be such that $\overline{f_i}=f_i$. Then $\bar{\varepsilon}_i=\varepsilon_i$ and $\beta_i=\max(\varepsilon_i,\,p^s-\varepsilon_i)$. Since $\max(\varepsilon_i,\,p^s-\varepsilon_i)\geqslant p^s/2$ for all possible values of ε_i, $\Pr\{\beta_i=j\}=0$ for all j, $j<p^s/2$.
  If $j>p^s/2$, then $j\neq p^s-j$, so

  $$\Pr\{\beta_i=j\}=\Pr\{\varepsilon_i=j\}+\Pr\{\varepsilon_i=p^s-j\}=2/(p^s+1).$$

  If p^s is even and $j=p^s/2$, then

  $$\Pr\{\beta_i=p^s/2\}=\Pr\{\varepsilon_i=p^s/2\}=1/(p^s+1).$$

  Putting these results in the formula $E(\beta_i)=\sum_{j=0}^{p^s}\Pr\{\beta_i=j\}j$, we get $E(\beta_i)=(3p^s+1)/4$, if p^s is odd, and $E(\beta_i)=(3p^s+1)/4-1/(4(p^s+1))$, otherwise. Putting both cases together, we get

  $$E(\beta_i)=\frac{3p^s+1}{4}-\frac{\delta_{p^s}}{4(p^s+1)}.$$

- Let i be such that $\overline{f_i}\neq f_i$. Then ε_i and $\bar{\varepsilon}_i$ are independent. We get

  $$\begin{aligned}\Pr\{\beta_i=j\}&=\Pr\{(\varepsilon_i=j\wedge p^s-\bar{\varepsilon}_i\leqslant j)\vee(p^s-\bar{\varepsilon}_i=j\wedge\varepsilon_i\leqslant j)\}\\&=\Pr\{\varepsilon_i=j\}\Pr\{\bar{\varepsilon}_i\geqslant p^s-j\}\\&\quad+\Pr\{\bar{\varepsilon}_i=p^s-j\}\Pr\{\varepsilon_i\leqslant j\}\\&\quad-\Pr\{\varepsilon_i=j\}\Pr\{\bar{\varepsilon}_i=p^s-j\},\end{aligned}$$

  where "$\wedge$" and "$\vee$" mean "and" and "or", respectively. Using the equality $\Pr\{\varepsilon_i\leqslant j\}=\sum_{l=0}^{j}\Pr\{\varepsilon_i=l\}$, analogous formula for $\Pr\{\bar{\varepsilon}_i\geqslant p^s-j\}$, and known probabilities of ε_i, we get $\Pr\{\beta_i=j\}=(2j+1)/(p^s+1)^2$ and

  $$E(\beta_i)=\frac{p^s(4p^s+5)}{6(p^s+1)}.$$

So we get

$$\begin{aligned} E_q(n) &= n - \sum_{i=1}^{t} (\deg f_i) E(\beta_i) \\ &= n - \sum_{i\in J} (\deg f_i) E(\beta_i) - \sum_{i\notin J} (\deg f_i) E(\beta_i) \\ &= n - B_{r,q}\left(\frac{3p^s+1}{4} - \frac{\delta_{p^s}}{4(p^s+1)}\right) - (r - B_{r,q})\frac{p^s(4p^s+5)}{6(p^s+1)} \\ &= n\left(\frac{1}{3} - \frac{1}{6(p^s+1)}\right) - B_{r,q}\left(\frac{p^s+1}{12} + \frac{2-3\delta_{p^s}}{12(p^s+1)}\right). \qquad \square \end{aligned}$$

Proposition 23. $B_{r,q} = \sum_{l|r,\, l\in\mathcal{N}_q} \varphi(l)$.

Proof. By Propositions 17 and 15, the definition of the cyclotomic polynomial $Q_d(X)$, and Proposition 16, we have

$$\begin{aligned} B_{r,q} &= \sum_{i\in J} \deg f_i = \sum_{i:\mathrm{ord}(f_i)\in\mathcal{N}_q} \deg f_i = \sum_{l|r,\, l\in\mathcal{N}_q} \sum_{i:\mathrm{ord}(f_i)=l} \deg f_i \\ &= \sum_{l|r,\, l\in\mathcal{N}_q} \deg Q_l(X) = \sum_{l|r,\, l\in\mathcal{N}_q} \varphi(l). \qquad \square \end{aligned}$$

We simplify further the expression of $B_{r,q}$ using the $\mathcal{N}_q$-factorization of r.

Proposition 24. *Let $r = 2^\beta d' d_1 d_2 d_3 \cdots$ be the $\mathcal{N}_q$-factorization of r. Let $\gamma \geqslant 0$ be such that $2^\gamma \,\|\, q+1$. Then $B_{r,q} = 2^{\min(\gamma,\beta)} d_1 + 2^{\min(\beta,1)} \sum_{\alpha\geqslant 2} (d_\alpha - 1)$.*

Proof. By Proposition 23 $B_{r,q} = \sum_{l|r,\, l\in\mathcal{N}_q} \varphi(l)$. If a divisor l of r belongs to $\mathcal{N}_q$, it belongs to one of $\mathscr{P}_{q,\alpha}$, $\alpha \geqslant 0$, so it divides $2^{\min(\beta,1)}$, if $\alpha = 0$, $2^{\min(\gamma,\beta)} d_1$, if $\alpha = 1$, or $2^{\min(\beta,1)} d_\alpha$, if $\alpha \geqslant 2$ (this follows from the definition of $\mathcal{N}_q$-factorization, Corollary 5 and the fact that l is a divisor of r).

Suppose that $\beta = 0$. Due to the partitioning of $\mathcal{N}_q$ into $\{\mathscr{P}_{q,\alpha}\}_{\alpha\geqslant 0}$ sets, each divisor $l > 1$ of r belonging to $\mathcal{N}_q$ divides exactly one of above-mentioned divisors of r, so using the formula $\sum_{i|j} \varphi(i) = j$ we get

$$\begin{aligned} B_{r,q} &= \sum_{l|r,\, l\in\mathcal{N}_q} \varphi(l) = \varphi(1) + \sum_{\alpha\geqslant 1} \sum_{d|d_\alpha,\, d\neq 1} \varphi(d) \\ &= 1 + \sum_{\alpha\geqslant 1} (d_\alpha - 1) = d_1 + \sum_{\alpha\geqslant 2} (d_\alpha - 1). \end{aligned}$$

When $\beta \geqslant 1$, in the same way we get that

$$\begin{aligned} B_{r,q} &= \sum_{l|r,\, l\in\mathcal{N}_q} \varphi(l) \\ &= \varphi(1)+\varphi(2)+\sum_{d|2^{\min(\gamma,\beta)}d_1,\, d\neq 1,\, d\neq 2} \varphi(d)+\sum_{\alpha\geqslant 2}\sum_{d|2d_\alpha,\, d\neq 1,\, d\neq 2}\varphi(d) \\ &= 1+1+2^{\min(\gamma,\beta)}d_1-2+\sum_{\alpha\geqslant 2}(2d_\alpha-2) \\ &= 2^{\min(\gamma,\beta)}d_1+2\sum_{\alpha\geqslant 2}(d_\alpha-1). \end{aligned}$$

Both cases sum up to the desired formula. □

Proof of Theorem 10. It is a direct consequence of Propositions 22 and 24. □

4. Roots and bounds

Theorem 25. *Let q be a power of a prime p, let $n \geqslant 1$. Then*:

(1) $E_q(n) = 0$ *if, and only if,* $n \in \mathcal{N}_q$.
(2) $n/12 \leqslant E_q(n) < n/3$ *for all,* $n \notin \mathcal{N}_q$.

Proof. Let as above $n = rp^s$, $s \geqslant 0$, $(r, p) = 1$. We begin by proving the first part of the theorem. Using the expression of $E_q(n)$ of Theorem 10, we get that $E_q(n) = 0$ if, and only if,

$$\frac{B_{r,q}}{r} = \frac{4p^{2s}+2p^s}{p^{2s}+2p^s+3-3\delta_{p^s}}. \tag{8}$$

It is not difficult to show that the right-hand side is $\geqslant 1$, with equality if, and only if, $p^s = 1$. But by definition of $B_{r,q}$ (Eq. (7)), we have $B_{r,q} \leqslant r$, so Eq. (8) holds if, and only if, $p^s = 1$ and $B_{r,q} = r$. But $B_{r,q} = r$ if, and only if, $r \in \mathcal{N}_q$. Indeed, $B_{r,q} = r$ signifies that all f_i from Eq. (3) are qsr, that is, for all f_i, $\mathrm{ord}(f_i) \in \mathcal{N}_q$ (Proposition 17). In particular, $r \in \mathcal{N}_q$. Conversely, if $r \in \mathcal{N}_q$, then by part 2 of Lemma 1 any positive divisor d of r belongs to $\mathcal{N}_q$, so that $\mathrm{ord}(f_i) \in \mathcal{N}_q$ for all f_i and $B_{r,q} = r$. Thus we get that Eq. (8) holds if, and only if, $p^s = 1$ and $r \in \mathcal{N}_q$. By part 1 of Lemma 1, these two properties sum up to $n \in \mathcal{N}_q$. So the first part of the theorem is proved.

Now the second part. Part 1 of Corollary 11 shows that $E_q(n) < n/3$ for all $n \geqslant 1$. It remains to prove that $E_q(n) \geqslant n/12$ for all $n \geqslant 1$, $n \notin \mathcal{N}_q$. Suppose that q and n are not relatively prime, so $p^s \geqslant 2$. Remark that Eq. (1) and part 2 of Corollary 11 involve the same quantity $B_{r,q}$. This allows us to express $E_q(n)$ as a function of $E_q(r)$. We get

$$E_q(n) = \frac{n}{4} - \frac{r}{4} + \frac{r\delta_{p^s}}{4(p^s+1)} + E_q(r)\left(\frac{p^s+1}{3} + \frac{2-3\delta_{p^s}}{3(p^s+1)}\right).$$

Using this expression, it is not difficult to show that

$$\frac{E_q(n)}{n} \geqslant \frac{1}{4} - \frac{1}{4p^s} + \frac{\delta_{p^s}}{4p^s(p^s+1)}.$$

If $p^s = 2$, then $E_q(n)/n \geqslant \frac{1}{4} - \frac{1}{8} + \frac{1}{24} = \frac{1}{6}$. If $p^s \geqslant 3$, then $E_q(n)/n \geqslant \frac{1}{4} - 1/(4p^s) \geqslant \frac{1}{4} - \frac{1}{12} = \frac{1}{6}$. Therefore, $E_q(n) \geqslant n/6$, when q and n are not relatively prime.

Suppose now that q and n are relatively prime. Since in this case $E_q(n) = (n - B_{n,q})/4$, we must show that $B_{n,q} \leqslant 2n/3$ for all $n \geqslant 1$, $n \notin \mathcal{N}_q$. Let $n \notin \mathcal{N}_q$. Let $n = 2^\beta d' d_1 d_2 d_3 \cdots$ be the $\mathcal{N}_q$-factorization of n. By Proposition 23

$$\frac{B_{n,q}}{n} = \frac{2^{\min(\gamma,\beta)} d_1 + 2^{\min(\beta,1)} \sum_{\alpha \geqslant 2} (d_\alpha - 1)}{2^\beta d' d_1 d_2 d_3 \cdots}.$$

Let $\alpha_1 < \alpha_2 < \cdots < \alpha_j$, where $j \geqslant 0$, be the indices α such that $d_\alpha > 1$. Note that $(d_\alpha > 1 \Rightarrow d_\alpha \geqslant 3)$ for all $\alpha \geqslant 1$ (this is valid also for d'). We consider six cases: (1) $d' = 1$ and $j = 0$, (2) $d' = 1$ and $j = 1$, (3) $d' > 1$ and $j = 0$, (4) $d' > 1$ and $j = 1$, (5) $j = 2$, (6) $j \geqslant 3$.

(1) By Lemma 9 $\beta \geqslant \gamma + 1$, so that $B_{n,q}/n = \frac{1}{2^{\beta-\gamma}} \leqslant \frac{1}{2}$.

(2) If $\alpha_1 = 1$, then by the same Lemma $\beta \geqslant \gamma + 1$, so we get $B_{n,q}/n \leqslant \frac{1}{2}$ in the same way. Otherwise, by the same Lemma $\beta \geqslant 2$, so

$$\frac{B_{n,q}}{n} = \frac{2^{\min(\gamma,\beta)}}{2^\beta d_{\alpha_1}} + \frac{2d_{\alpha_1}}{2^\beta d_{\alpha_1}} - \frac{2}{2^\beta d_{\alpha_1}} \leqslant \frac{1}{d_{\alpha_1}} + \frac{1}{2^{\beta-1}} - \frac{1}{2^{\beta-1}} \frac{1}{d_{\alpha_1}}.$$

It is not difficult to show that the function $x+y-xy$, where $0 < x \leqslant \frac{1}{3}$, $0 < y \leqslant \frac{1}{2}$, reaches its maximum when $x = \frac{1}{3}$, $y = \frac{1}{2}$, so that the last expression is $\leqslant \frac{1}{3} + \frac{1}{2} - \frac{1}{2} \cdot \frac{1}{3} = \frac{2}{3}$.

Now let us replace $B_{n,q}/n$ by a simpler expression:

$$\frac{B_{n,q}}{n} \leqslant \frac{d_1 + \sum_{\alpha \geqslant 2} (d_\alpha - 1)}{d' d_1 d_2 d_3 \cdots} = \frac{1 - j + \sum_{i=1}^{j} d_{\alpha_i}}{d' d_{\alpha_1} d_{\alpha_2} \cdots d_{\alpha_j}}.$$

(3) $B_{n,q}/n \leqslant 1/d' \leqslant \frac{1}{3}$.

(4) $B_{n,q}/n \leqslant d_{\alpha_1}/(d' d_{\alpha_1}) \leqslant \frac{1}{3}$.

As $j \geqslant 2$ and $d' \geqslant 1$ for two remaining cases, we may replace $B_{n,q}/n$ by even simpler expression. Let i_0, $1 \leqslant i_0 \leqslant j$, be the index such that $d_{\alpha_{i_0}} = \max_{1 \leqslant i \leqslant j} d_{\alpha_i}$. Then

$$\frac{B_{n,q}}{n} \leqslant \frac{\sum_{i=1}^{j} d_{\alpha_i}}{\prod_{i=1}^{j} d_{\alpha_i}} \leqslant \frac{j d_{\alpha_{i_0}}}{\prod_{1 \leqslant i \leqslant j} d_{\alpha_i}} = \frac{j}{\prod_{1 \leqslant i \leqslant j,\, i \neq i_0} d_{\alpha_i}}.$$

(5) The set of indices is $\{\alpha_1, \alpha_2\}$, so i_0 is 1 or 2. Let i_1 be the other index. Then $B_{n,q}/n \leqslant 2/d_{\alpha_{i_1}} \leqslant 2/3$.

(6) Let j_0, $j-1 \leqslant j_0 \leqslant j$, be an index such that $j_0 \neq i_0$. Then $j < 2^{j-1} \leqslant 2^{j_0} \leqslant 2^{\alpha_{j_0}} < d_{\alpha_{j_0}}$ by Lemma 8, so

$$\frac{B_{n,q}}{n} < \frac{1}{\prod_{1 \leqslant i \leqslant j,\ i \neq i_0,\ i \neq j_0} d_{\alpha_i}} \leqslant \frac{1}{3}.$$

In all six cases $B_{n,q} \leqslant 2n/3$. □

5. Asymptotic behavior of $E_q(n)$

In this section we study the asymptotic behavior of $E_q(n)/n$, using Theorem 10. We already know (Theorem 25) that either $E_q(n) = 0$ (if, and only if, $n \in \mathscr{N}_q$), or $\frac{1}{12} \leqslant E_q(n)/n < \frac{1}{3}$ for all $n \geqslant 1$, $n \notin \mathscr{N}_q$. In this section we consider the sequence

$$\left(\frac{E_q(n)}{n}\right)_{n \geqslant 1,\, n \notin \mathscr{N}_q}. \tag{9}$$

We find its upper and lower limits. Moreover, we find the set of limits of all converging subsequences of the sequence

$$\left(\frac{E_q(n)}{n}\right)_{n \geqslant 1,\,(n,q)=1} \tag{10}$$

(these limits can be accumulation points or isolated points reached an infinity of times, i.e. values taken by stationary subsequences).

Theorem 26. *The set of limits of all converging subsequences of sequence* (10) *is*

$$D = \left\{\frac{1}{4}\right\} \cup \left\{\frac{1}{4}\left(1 - \frac{1}{k}\right) \middle| \, k \geqslant 1,\ (k,q) = 1\right\}.$$

Proof. Note that, since n and q are relatively prime, by Theorem 10 and Corollary 11

$$\frac{E_q(n)}{n} = \frac{1}{4}\left(1 - \frac{B_{n,q}}{n}\right),$$

where

$$\frac{B_{n,q}}{n} = \frac{2^{\min(\gamma,\beta)} d_1 + 2^{\min(\beta,1)} \sum_{\alpha \geqslant 2} (d_\alpha - 1)}{2^\beta d' d_1 d_2 d_3 \cdots},$$

$n = 2^\beta d' d_1 d_2 d_3 \cdots$ is the $\mathscr{N}_q$-factorization of n, γ is defined by $2^\gamma \,\|\, q+1$.

Note also that if $t \in \mathscr{P}_{q,\alpha}$, $\alpha \geqslant 1$, t odd, then the $\mathscr{N}_q$-factorization of nt differs very little from the one of n, more precisely, $\beta(nt) = \beta(n)$, $d'(nt) = d'(n)$, $d_i(nt) = d_i(n)$ for all $i \geqslant 1$, $i \neq \alpha$, and the only difference is the value of d_α: $d_\alpha(nt) = t d_\alpha(n)$. When $t \in \mathscr{P}'_q$, the only difference is the value of d': $d'(nt) = t d'(n)$. And if $t = 2^\eta$ and q is odd, the only difference is the value of β: $\beta(nt) = \beta(n) + \eta$ (see Section 2 for the definitions of $\beta(n)$, $d'(n)$, $d_i(n)$, $\mathscr{P}_{q,\alpha}$ and $\mathscr{P}'_q$).

We begin by proving that for any point $\xi \in D$, there exists a subsequence of sequence (10) tending to ξ as $n \to \infty$. Let $\xi \in D$, i.e. $\xi = \frac{1}{4}$ or $\xi = (1-1/k)/4,\ k \geqslant 1,\ (k,q) = 1$. Let first $\xi = (1-1/k)/4,\ k \geqslant 1,\ (k,q) = 1$. Let $j \geqslant 2$ be such that $d_j(k) = 1$.

Assume that $\beta(k) = 0$. Then for $t \in \mathscr{P}_{q,j},\ t$ odd,

$$\frac{B_{kt,q}}{kt} = \frac{2^{\min(\gamma,\beta(k))} d_1(k) + 2^{\min(\beta(k),1)} \sum_{\alpha \geqslant 2,\ \alpha \neq j} (d_\alpha(k) - 1)}{kt} + \frac{2^{\min(\beta(k),1)} (d_j(k)t - 1)}{kt},$$

so $B_{kt,q}/kt$ tends to $2^{\min(\beta(k),1)} d_j(k)/k = 1/k$ as t tends to infinity, running through the odd numbers of $\mathscr{P}_{q,j}$. Thus, $E_q(kt)/kt$ tends to ξ.

Now assume that $\beta(k) \geqslant 1$. In the same way we get that $B_{2kt,q}/2kt$ tends to $2^{\min(\beta(k)+1,1)} d_j(k)/2k = 2d_j(k)/2k = 1/k$ as $t \to \infty,\ t \in \mathscr{P}_{q,j},\ t$ odd, thus $E_q(2kt)/2kt$ tends to ξ. Note that this is indeed a subsequence of sequence (10), since $\beta(k) \geqslant 1$ and $(k,q) = 1$ imply that q is odd, so $(2kt, q) = 1$.

It remains the case when $\xi = \frac{1}{4}$. Let $\alpha_1 \geqslant 1,\ \alpha_2 \geqslant 1,\ \alpha_1 \neq \alpha_2$. In the same way we get that $B_{t_1 t_2,q}/t_1 t_2$ tends to 0 as $t_1 \to \infty,\ t_2 \to \infty,\ t_1 \in \mathscr{P}_{q,\alpha_1},\ t_2 \in \mathscr{P}_{q,\alpha_2},\ t_1$ and t_2 are odd. Thus, $E_q(t_1 t_2)/t_1 t_2$ tends to $\xi = \frac{1}{4}$.

Now we prove that if a subsequence of sequence (10) converges, its limit belongs to D. Let

$$\left(\frac{E_q(n)}{n}\right)_{n \in J} \tag{11}$$

be a converging subsequence of sequence (10). If it converges to $\frac{1}{4}$, then the theorem is proved. Suppose that it converges to a number different from $\frac{1}{4}$, that is, $(B_{n,q}/n)_{n \in J}$ converges to a non-zero number w. The sequence (11) defines the sequences $(d_\alpha(n))_{n \in J}$ for all $\alpha \geqslant 1$. If all these sequences $(d_\alpha(n))_{n \in J},\ \alpha \geqslant 1$, are bounded, then $B_{n,q}$ is bounded, so $B_{n,q}/n$ tends to zero as $n \to \infty,\ n \in J$, a contradiction. So there exists at least one non-bounded sequence, say, the sequence $(d_i(n))_{n \in J}$ for one $i \geqslant 2$ (if $i = 1$, the proof is analogous, it suffices to replace $\min(\beta(n), 1)$ by $\min(\beta(n), \gamma)$). Then there exists a subsequence $(d_i(n))_{n \in K},\ K \subset J$, of the sequence $(d_i(n))_{n \in J}$, tending to infinity. The sequence $(B_{n,q}/n)_{n \in J}$ converges to $w \neq 0$, so its subsequence $(B_{n,q}/n)_{n \in K}$ converges to $w \neq 0$ too. All terms of $B_{n,q}/n$ tends to zero as $n \to \infty,\ n \in K$, except probably the only term not having $d_i(n)$ in its denominator, that is, the term

$$\frac{2^{\min(\beta(n),1)} d_i(n)}{n} = \frac{1}{2^{\max(\beta(n)-1,0)} d'(n) d_1(n) \cdots d_{i-1}(n) d_{i+1}(n) \cdots},$$

which therefore must tend to $w \neq 0$. So the sequence

$$(2^{\max(\beta(n)-1,0)} d'(n) d_1(n) \cdots d_{i-1}(n) d_{i+1}(n) \cdots)_{n \in K}$$

converges to $1/w$. But it is a sequence of integers, so it stabilizes, it becomes equal to an integer $k = 1/w \geqslant 1$. This integer k is relatively prime to q, since n runs through integers

relatively prime to q. So the sequence $(B_{n,q}/n)_{n\in J}$ converges to $1/k$, $k \geqslant 1$, $(k,q)=1$, and sequence (11) converges to an element of the set D. □

So we proved that the set of limits of all converging subsequences of sequence (10) is the set D. The study of the set of limits of all converging subsequences of sequence (9) is much more complicated, and we do not investigate it here.

Theorem 27. *The upper limit of sequence* (9) *is* $\frac{1}{3}$; *its lower limit is* $\frac{1}{6}$, if $p=2$, *and* $\frac{1}{8}$ *otherwise. The upper limit of the sequence*

$$\left(\frac{E_q(n)}{n}\right)_{n\geqslant 1,\, n\notin \mathscr{N}_q,\ (n,q)=1} \tag{12}$$

is $\frac{1}{4}$; *its lower limit is equal to the one of sequence* (9).

Proof. Sequence (12) is a subsequence of sequence (10), so the set of limits of all converging subsequences of sequence (12), denoted by Lim(12), is a subset of Lim(10). The difference between sequences (10) and (12) is the values of $E_q(n)/n$ for $n\in\mathscr{N}_q$. But all these values are equal to zero (part 1 of Theorem 25). So the only difference between Lim(12) and Lim(10) may be the point 0. And indeed $0\notin$ Lim(12), since by part 2 of Theorem 25 $E_q(n)/n \geqslant \frac{1}{12}$ for all $n\geqslant 1$, $n\notin\mathscr{N}_q$, so there exists no subsequence of sequence (12) tending to zero. Thus, $\mathrm{Lim}(12)=\mathrm{Lim}(10)\setminus\{0\}=\{\frac{1}{4}\}\cup\{(1-1/k)/4 \mid k\geqslant 2,(k,q)=1\}$ (Theorem 26).

The lower and upper limits of sequence (12) are, respectively, the smallest and largest elements of Lim(12). So the upper limit is $\frac{1}{4}$, and the lower limit is $(1-\frac{1}{3})/4=\frac{1}{6}$, if $2\mid q$, and $(1-\frac{1}{2})/4=\frac{1}{8}$, otherwise.

In the proof of Theorem 25 we have seen that $E_q(n)/n\geqslant\frac{1}{6}$ for all $n\geqslant 1$ which are not relatively prime to q. Therefore, for all $n\geqslant 1$ which are not relatively prime to q, $E_q(n)/n$ is greater than or equal to the lower limit of sequence (12), so the lower limit of sequence (9) is equal to the one of sequence (12).

By part 1 of Corollary 11, $E_q(n)/n<\frac{1}{3}$ for all $n\geqslant 1$, $n\notin\mathscr{N}_q$. Moreover, there exists a subsequence of sequence (9) converging to $\frac{1}{3}$. Indeed, let $\alpha_1\geqslant 1$, $\alpha_2\geqslant 1$, $\alpha_1\neq\alpha_2$, $t_1\in\mathscr{P}_{q,\alpha_1}$, $t_2\in\mathscr{P}_{q,\alpha_2}$, t_1 and t_2 odd, $s\geqslant 1$. Then $(E_q(t_1t_2p^s)/t_1t_2p^s)$ is a subsequence of sequence (9), since $t_1t_2p^s\notin\mathscr{N}_q$ follows from $s\geqslant 1$ (part 1 of Lemma 1). And, by Theorem 10

$$\frac{E_q(t_1t_2p^s)}{t_1t_2p^s}=\frac{1}{3}-\frac{1}{6(p^s+1)}-\frac{B_{t_1t_2,q}}{t_1t_2}\left(\frac{p^s+1}{12p^s}+\frac{2-3\delta_{p^s}}{12p^s(p^s+1)}\right).$$

In the proof of Theorem 26 we saw that $B_{t_1t_2,q}/t_1t_2$ tends to 0 as $t_1\to\infty$, $t_2\to\infty$, $t_1\in\mathscr{P}_{q,\alpha_1}$, $t_2\in\mathscr{P}_{q,\alpha_2}$, t_1 and t_2 odd. Thus, $E_q(t_1t_2p^s)/t_1t_2p^s$ tends to $\frac{1}{3}$ as $s\to\infty$, $t_1\to\infty$, $t_2\to\infty$, $t_1\in\mathscr{P}_{q,\alpha_1}$, $t_2\in\mathscr{P}_{q,\alpha_2}$, t_1 and t_2 odd. Thus the upper limit of sequence (9) is $\frac{1}{3}$. □

Table 1
Example of the values of $\mathcal{N}_q(x)$

q	2	3	4	5	7	8	9	11
$\mathcal{N}_q(10^4)$	1523	1876	791	2150	1849	1523	776	2154
$\mathcal{N}_q(10^6)$	110837	133493	57873	152929	132572	110837	52005	154955

6. Conclusions

In this paper we have studied $E_q(n)$, the average dimension of the hull of cyclic codes of length n over the finite field $\mathbf{F}_q$. We have derived an expression of $E_q(n)$ which handles well (Theorem 10). Using this expression, we have proved that either $E_q(n)$ is zero (if, and only if, $n \in \mathcal{N}_q$), or it grows at the same rate as n, when $n \notin \mathcal{N}_q$ (Theorem 25). Moreover, we have determined the set of limits of all converging subsequences of the sequence $(E_q(n)/n)_{n\geqslant 1,\ (n,q)=1}$, and we have found the upper and lower limits of the sequence $(E_q(n)/n)_{n\geqslant 1,\ n\notin\mathcal{N}_q}$.

It remains to discuss the size of $\mathcal{N}_q$. Write $\mathcal{N}_q(x)$ for the number of elements in the set $\mathcal{N}_q$ which are smaller than or equal to x. Dubickas et al., proved [2] the following result:

Theorem 28. *Given an integer $q \geqslant 2$, there is a positive constant c such that*

$$\mathcal{N}_q(x) < c\,\frac{x}{(\log x)^{1/4}}.$$

More precisely, $\mathcal{N}_q(x) < cx/(\log x)^{1/2}$ for q a square, and for every $\varepsilon > 0$ there is a positive constant $c = c(\varepsilon)$ such that $\mathcal{N}_2(x) < cx/(\log x)^{7/24-\varepsilon}$, and $\mathcal{N}_q(x) < cx/(\log x)^{1/3-\varepsilon}$ for $q \geqslant 3$ square-free.

Corollary 29. *Almost all positive integers do not belong to $\mathcal{N}_q$, i.e. $\mathcal{N}_q(x)/x \to 0$ as $x \to \infty$.*

But, on the other hand, the set $\mathcal{N}_q$ is not so small. For example, for $q = 2$ or $q \geqslant 3$ square-free, there is a constant c such that $\mathcal{N}_q(x) > cx/\log x$ [2]. In Table 1 we give the values $\mathcal{N}_q(10^4)$ and $\mathcal{N}_q(10^6)$ for first several values of q.

As mentioned in the introduction, the algorithm presented in [11–13] is limited by the size of the hull. The results of the present paper show that there exist "not many" $n \geqslant 1$ such that the hull of most cyclic codes of length n is "small". More precisely, if a is a positive constant, then for almost all $n \geqslant 1$, $E_q(n) > a$. Indeed, $E_q(n) \geqslant n/12$ for almost all $n \geqslant 1$ (Theorem 25 and Corollary 29), therefore $E_q(n) > a$ for almost all $n > 12a$, thus for almost all $n \geqslant 1$. As a conclusion, for almost all $n \geqslant 1$, most of cyclic codes of length n are not accessible to this algorithm.

Acknowledgements

I should like to thank one of the reviewers for his remarks and suggestions about Proposition 13 and Theorem 25, and both reviewers for their assistance in improving the readability of the text.

References

[1] E.F. Assmus Jr., J.D. Key, Affine and projective planes, Discrete Math. 83 (1990) 161–187.
[2] A. Dubickas, F. Holland, G. Skersys, On integers dividing shifted powers and their applications to coding theory, 2002, in preparation.
[3] D. Knee, H.D. Goldman, Quasi-self-reciprocal polynomials and potentially large minimum distance BCH codes, IEEE Trans. Inform. Theory IT-15 (1) (1969) 118–121.
[4] J. Leon, Computing automorphism groups of error-correcting codes, IEEE Trans. Inform. Theory IT-28 (3) (1982) 496–511.
[5] J. Leon, Permutation group algorithms based on partitions, I: theory and algorithms, J. Symbolic Comput. 12 (1991) 533–583.
[6] J. Leon, Partitions, refinements, and permutation group computation, DIMACS, Discrete Math. Theoret. Comput. Sci. 28 (1997) 123–158.
[7] R. Lidl, H. Niederreiter, Finite Fields, Cambridge University Press, Cambridge, 1997.
[8] F.J. MacWilliams, N.J.A. Sloane, The Theory of Error-Correcting Codes, North-Holland, Amsterdam, 1978.
[9] N. Sendrier, On the dimension of the hull, SIAM J. Appl. Math. 10 (1997) 282–293.
[10] N. Sendrier, Finding the permutation between equivalent codes: the support splitting algorithm, IEEE Trans. Inform. Theory IT-46 (4) (2000) 1193–1203.
[11] N. Sendrier, G. Skersys, Permutation groups of error-correcting codes, in: D. Augot, C. Carlet (Eds.), Proceedings of Workshop on Coding and Cryptography, INRIA, Paris, 1999, pp. 33–41.
[12] N. Sendrier, G. Skersys, On the computation of the automorphism group of a linear code, in: Proceedings of IEEE ISIT'2001, Washington, DC, 2001, p. 13.
[13] G. Skersys, Calcul du groupe d'automorphismes des codes. Détermination de l'équivalence des codes, Ph.D. Thesis, Limoges University, Limoges, 1999.